Scalable Video-on-Demand Streaming for Heterogeneous Clients in Wireless Network

Li Zhou [1], Xiaohua Tian [1,2], Xiaoying Gan [1], Hui Yu [1], Xinbing Wang [1]
[1] Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2] National Mobile Communications Research Laboratory, Southeast University, China
Email: {sjtu_julian, xtian, ganxiaoying, yuhui, xwang8}@sjtu.edu.cn

***Abstract*—Periodic broadcasting has achieved prominent performance in VoD (Video on Demand) service in wired network. However, the development of wireless VoD service still has gone on hard and slowly in comparison to the rapid growth of mobile video service. In this paper, we propose a scalable video streaming method HQOBA (Heterogeneous Quality-Oriented Bandwidth Allocation) to address the problems in wireless network. To be concrete, in order to tackle the heterogeneous client type problem, we employ variable bandwidth allocation, instead of fixed bandwidth value for each channel. Moreover, to handle the problem of scarce server bandwidth and unpredictable wireless channel conditions, the proposed method takes the advantage of FGS (Fine Granular Scalable) video coding and Periodic Broadcasting to successfully transmit video segment. Furthermore, we develop a bandwidth allocation algorithm to maximize client perceptual visual quality for the proposed transmit scheme. The experimental results of HQOBA scheme also validate the effectiveness.**

***Index Terms*—Wireless video-on-demand, bandwidth allocation, heterogeneous clients.**

I. Introduction

Mobile video service has been experiencing an explosive growth with the rapid advancement in wireless network. Considering the forecast that 75 percent of the mobile data traffic in the world will be consumed by video by 2020 [1], and wireless VoD (Video-on-Demand) service may become killer application in the future. VoD system has two features: 1) 5% to 10% popular video contents constitute most video request [2]. 2) it may easily run out of server bandwidth because the growth of bandwidth can never keep up with the growth of the client number [3]. It is acknowledged that the bandwidth of server will not change with the increase of the client number when applying periodic broadcasting, making it a suitable choice to distribute popular video contents in wired network context. Many periodic broadcasting protocols have been proposed to save the server bandwidth in wired network such as Pyramid Broadcasting [4], Fast Broadcasting [4], Harmonic Broadcasting [4], CCA [5], CCA+ [6], CAR [7] and DeRe [8].

However, transplanting the periodic broadcasting scheme to wireless VoD systems is still difficult, the challenge lies in three aspects. First, the bandwidth of server is scarce for a VoD system in wireless network compared with wired network. When the number of videos to be broadcasted is large, the server may not be capable of distributing all the entire video contents simultaneously. Second, unlike the stable wired channel, the wireless channel condition is unpredictable. It may incur packet loss or even discontinuous playing of requested video. Third, considering the heterogeneity of client reception bandwidth, the wireless video streaming scheme should be scalable to meet the perceptual quality demand of different clients based on their reception bandwidth.

QABA [9] allows scalable transmission of requested media by FGS video coding [10]. Nevertheless, it assumes that all clients are capable of receiving the complete video all the time, which can be hardly implemented in wireless network because of the heterogeneity of client reception bandwidth. For example, many mobile video applications, like *Youtube* and *Youku*, allow clients to choose different levels of video quality for continuous playing. In a word, QABA just focuses on the server bandwidth scarcity and wireless channel condition without the heterogeneity of client reception bandwidth. Although OHPB [11] provides a scalable periodic broadcasting protocol for heterogeneous clients, it massively increases the complexity of channel management due to the linear growth of the number of broadcasting channels. To achieve the same quality as our proposed method with same number of channels, it will cost a much larger minimum access latency.

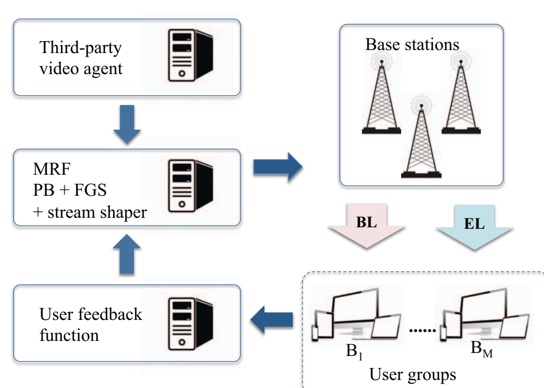
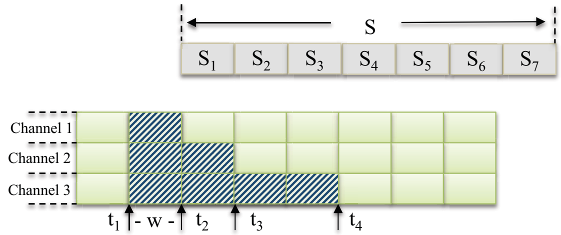
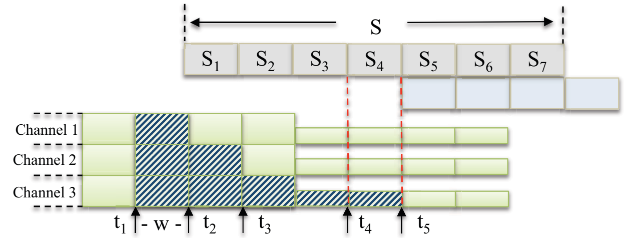


Fig. 1. System model of scalable VoD streaming for heterogeneous clients.

This work is supported by National Natural Science Foundation of China (No. 61572319, U1405251, 61532012, 61325012, 61271219, 61428205); National Mobile Communications Research Laboratory, Southeast University (No. 2014D07).



(a) Continuous playout with stable channel.



(b) Incontinuous playout with time-varing channel.

Fig. 2. Fast broadcasting examples with different channel condition.

In this paper, we propose an HQOBA (Heterogeneous Quality-Oriented Bandwidth Allocation) method to maximize the client perceptual quality. Specifically, we address the problem of heterogeneous client reception bandwidth with the employment of variable bandwidth allocation, instead of fixed bandwidth value for each channel. Moreover, periodic broadcasting is imposed to save server bandwidth. Finally, the model is improved by adding FGS video coding to accommodate the time-varying channel condition because of the great scalability of FGS to accommodate dynamic bit-rate of wireless channel.

Remarkably, the contribution of this paper is three-fold. First, the HQOBA supports heterogeneous client reception bandwidth better than previous wireless video streaming schemes. Second, it provides a scalable broadcasting scheme in wireless network with the combination of periodic broadcasting and FGS video coding. Third, a bandwidth allocation algorithm is developed to maximize client perceptual visual quality.

This paper is structured as follows. Section 2 presents background on periodic broadcasting and on quality adaptation for media streaming. This section also describes our basic system model. Section 3 outlines the HQOBA for heterogeneous clients followed by simulation results presented in section 4. Conclusions are presented in section 5.

II. System Model

The basic system model of the proposed method is illustrated in Fig. 1, which consists of video server agent, MRF (Media Resource Function), client feedback function and heterogeneous users. The media source data are placed in third-party video sever agent. The MRF is responsible for providing FGS video coding, periodic broadcasting protocol and adaptive stream shaping. Heterogeneous client bandwidth conditions are reported to client feedback function. Then the stream shaper can dynamically reshape the FGS encoded broadcasting stream based on the heterogeneous client bandwidth conditions and server bandwidth. The reshaped stream would be broadcasted to heterogeneous clients to maximize their perceptual visual quality.

In the following, we will briefly review periodic broadcasting protocols and different performances conditioned by variable channel context.

The crux of periodic broadcasting protocol is to fragment the video content into segments of different sizes and each segment is periodically broadcasted over different channels. In this way, VoD services can be provided in less server bandwidth cost with a given user access delay. Knowing that the proposed HQOBA can be easily extended to other periodic broadcasting protocols, we adopt FB (Fast Broadcasting) [12] protocol for the ease of exposition.

FB divides video into 2^{K-1} segments with the same sizes, and K is the total number of broadcast channels. A time slot $\frac{S}{2^{K-1}}$ is defined as the length of the time to broadcast a video segment, and S is the length of the whole video. In the example shown in Fig. 2 (a) , FB enables continuous playing with client access delay ω. This example is based on ideal channel condition, while wireless channel is much more unpredictable due to many reasons such as fading, shadowing, interference, path loss and noise. Fig. 2 (b) shows an example of FB with the deterioration of channel condition at t_3. Consequently, the discontinuous playing occurs between t_4 and t_5.

FGS video coding is used to guarantee continuous playing in wireless channel condition. A video stream in FGS is encoded into one BL (Base Layer) stream and one EL (Enhancement Layer) stream. The BL stream must be completely received to ensure basic quality received by clients and the EL stream is designed to enhance the visual quality. Note that the BL stream is typically encoded using a low bit-rate to accommodate a wide range of heterogeneous receivers, which provides clients of low perceptual quality. The BL stream is usually transmitted over QoS-guaranteed channels because of its importance and low bit-rate. The EL stream, which is encoded by a high bit-rate, can be decoded to enhance the visual quality based on the complete reception of BL stream. Accordingly, FGS provides better scalability because EL stream can be truncated at the level of bit granularity to achieve a desired target bit-rate [13].

Fig. 3 depicts the pipeline of the proposed method. First, FGS encodes video into BL stream and EL stream. Then, BL and EL are broadcasted separately with FB. After that, we get n equal-sized BL segments, say $s_b^1, s_b^2, ..., s_b^n$ BL stream and n equal-sized EL segments, say $s_e^1, s_e^2, ..., s_e^n$ from EL stream. A BL segment s_b^i and the corresponding EL segment s_e^i form a complete segment s_a^i. We assume that only BL streams are transmitted over bandwidth-guaranteed channels, while the EL

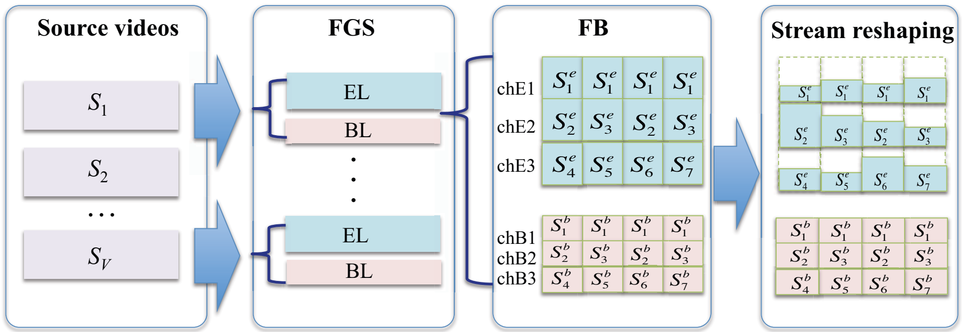


Fig. 3. Pipeline of the proposed method.

streams are designed to utilize the remaining server bandwidth. Hence, extending FB to FGS encoded video allows stream reshaping to save the bandwidth in wireless network.

III. Heterogeneous Quality-Oriented Bandwidth Allocation

We suppose that there are V videos broadcasted in our system with K BL channels and K EL channels. For each video v_i, it has k_i BL channels and k_i EL channels to transmit its FGS-encoded video content. Suppose there are M types of client reception bandwidth $[B_1, B_2, \cdots, B_M]$. We denote Q_t^{all} as the sum of entire stream visual quality received by all types of users during time slot t.

$$Q_t^{all} = Q_t^b + Q_t^e \tag{1}$$

where Q_t^b and Q_t^e represent the sum of stream visual quality received by all types of users during time slot t transmitted by BL and EL, Respectively. We discuss the situations under time slot t condition, hence we drop the subscript t for notational simplicity.

We denote the quality of the video segment client received by function $q(s)$. For video segment $s_i (i = 1, 2, \cdots, 2K)$, $q(s_i)$ can be calculated as [9] [14]

$$\forall i : q(s_i) = \beta_i \times x_i \tag{2}$$

where β_i is a constant coefficient and x_i is the bit-rate allocated to s_i.

Thus we use $q(s_i^b)$ and $q(s_i^e)$ to return visual quality of BL segment s_i^b and EL segment s_i^e with client reception bandwidth B_m^b and B_m^e, respectively. Since that $B_m = B_m^b + B_m^e$.

$$q(s_i^b, x_i^b) = \beta_i \times B_m^b \tag{3}$$

where x_i^b is the bit-rate allocated to s_i^b, x_i^e is the bit-rate allocated to s_i^e. Because of the low bit-rate of BL stream, we assume that the BL stream is transmitted over bandwidth-guaranteed channels. That is $B_m^b = x_i^{b,max}$. $x_i^{b,max}$ is the total bit-rate of BL encoded by FGS.

In contrast to $q(s_i^b, x_i^b)$, the visual quality of EL segment s_i^e in this paper is considered variably to adjust the bandwidth heterogeneity of different clients. We denote $q(s_i^e, x_i^e)$ as

$$\begin{aligned} q(s_i^e, x_i^e) &= \begin{cases} \beta_i \times x_i, & \text{if } B_m^e > x_i^e \\ \beta_i \times B_m^e, & \text{if } B_m^e \le x_i^e \end{cases} \\ &= \beta_i \times \min(B_m^e, x_i^e) \end{aligned} \tag{4}$$

To be concrete, if the client bandwidth is large enough to receive the EL segment transmitted by server, then $q(s_i^e, x_i^e)$ totally depends on the server terminal. However, if the client bandwidth is limited, which is a possible situation, then the first condition is not satisfied. In contrast, the B_m^e becomes the upper bound bandwidth the video segment can be received. Overall, the client perceptual visual quality provided by BL and EL segment at time slot t are expressed as

$$Q_t^b = \sum_{i=1}^{K} \sum_{m=1}^{M} q(s_i^b, x_i^b) \times \omega(s_i^b, B_m^b) \tag{5}$$

$$Q_t^e = \sum_{i=1}^{K} \sum_{m=1}^{M} q(s_i^e, x_i^e) \times \omega(s_i^e, B_m^e) \tag{6}$$

where ω represents the weight of each type of clients who receive a video segment.

As emphasized above that BL stream would be broadcasted over bandwidth-guaranteed channels, which means that the Q_t^b is a constant. In order to maximize Q_t^{all}, we just need to focus on how to allocate the remaining server bandwidth $X_{s,r}$ for EL channels. Then we can formulate our optimization as

$$\begin{aligned} \max \quad & \sum_{i=1}^{K} \sum_{m=1}^{M} q(s_i^e, x_i^e) \times \omega(s_i^e, B_m^e) \\ \text{s.t.} \quad & 0 \le x_i^e \le x_{i,max}^e \\ & \sum_{i=1}^{K} x_i^e = X_{s,r} \end{aligned} \tag{7}$$

where $x_{i,max}^e$ is the total bit-rate of EL encoded by FGS and $\omega(s_i^e, B_m^e)$ can be calculated as

$$\omega(s_i^e, B_m^e) = \overline{\mu}_m \times \alpha_i \times f(s_i^e) \tag{8}$$

where $\overline{\mu}_m$ denotes the mean number of type m arrived clients. α_i represents the access probability of the video containing segment s_i^e. $f(s_i^e)$ is the broadcasting probability of segment s_i^e.

Therefore equation (7) is equivalent to

$$\max \quad \sum_{i=1}^{K} \sum_{m=1}^{M} \beta_i \times \min(B_m^e, x_i^e) \times \overline{\mu}_m \times \alpha_i \times f(s_i^e)$$

$$\text{s.t.} \quad \begin{array}{l} 0 \leq x_i^e \leq x_{i,max}^e \\ \sum_{i=1}^{K} x_i^e = X_{s,r} \end{array}. \tag{9}$$

Then we will briefly prove that equation (9) is a convex optimization. First, we simplified it as

$$\max \quad \sum_{m=1}^{M} \sum_{i=1}^{K} \overline{\mu}_m \times P_i \times \min(B_m^e, x_i^e) \tag{10}$$

where $P_i = \beta_i \times \alpha_i \times f(s_i^e)$. Moreover, the non-linear function $\min(B_m^e, x_i^e)$ can be converted as

$$\min(B_m^e, x_i^e) = \frac{1}{2}(x_i^e + B_m^e) - \frac{1}{2}|x_i^e - B_m^e| \tag{11}$$

Thus the optimization problem (9) is equivalent to the following minimization problem

$$\min \quad \frac{1}{2} \sum_{m=1}^{M} \|\overline{\mu}_m \cdot \boldsymbol{P}\boldsymbol{E_{2,m}}\|_1 - \sum_{m=1}^{M} \overline{\mu}_m \cdot \boldsymbol{E_0}\boldsymbol{P}\boldsymbol{E_{1,m}} \tag{12}$$

where $\boldsymbol{E}_0 = \frac{1}{2}\boldsymbol{E}$. $\boldsymbol{E}$ is a unit vector, whose dimension is $1 \times K$. We denote $\boldsymbol{X} = (x_1^e, x_2^e, \cdots, x_K^e)^\top$, whose dimension is $K \times 1$. Accordingly, $\boldsymbol{E_{1,m}} = \boldsymbol{X} + B_m^e$ and $\boldsymbol{E_{2,m}} = \boldsymbol{X} - B_m^e$. $\boldsymbol{P}$ is diagonal matrix whose main diagonal element is P_i. The dimension of $\boldsymbol{P}$ is $K \times K$.

Given that L1-norm is convex [15], the first part of equation (12) which is a linear combination of L1-norm remains convex. As a linear function, the second part of equation (12) is convex as well as concave. Moreover, the constraint condition are two linear functions. Overall, we prove that equation (12) is a convex optimization problem, which can be solved by CVX package [16]. Meanwhile, we develop an allocation algorithm when server bandwidth is not sufficient, as algorithm 1.

Algorithm 1 HQOBA algorithm

Input: $\min \quad l(\boldsymbol{X}) + g(\boldsymbol{Z}) \quad \text{s.t.} \quad \boldsymbol{X} - \boldsymbol{Z} = 0$

$l(\boldsymbol{X}) = -\sum_{m=1}^{M} \overline{\mu}_m \cdot \boldsymbol{E_0}\boldsymbol{P}\boldsymbol{E_{1,m}} + \lambda(\boldsymbol{A}\boldsymbol{X} - X_{s,r})$

$g(\boldsymbol{Z}) = \frac{1}{2} \sum_{m=1}^{M} \|\overline{\mu}_m \cdot \boldsymbol{P}\boldsymbol{E_{2,m}}\|_1$

$Iteration \quad k \leftarrow 0$

repeat

$\boldsymbol{X}^{k+1} \leftarrow \mathrm{argmin}_{\boldsymbol{X}}(l(\boldsymbol{X}) + (\rho/2)\|\boldsymbol{X} - \boldsymbol{Z}^k + \boldsymbol{U}^k\|_2^2)$

$\boldsymbol{Z}^{k+1} \leftarrow S_{\lambda/\rho}(\boldsymbol{X}^{k+1} + \boldsymbol{U}^k)$

$\boldsymbol{U}^{k+1} \leftarrow \boldsymbol{U}^k + \boldsymbol{X}^{k+1} - \boldsymbol{Z}^{k+1}$

until converge criterion is satisified

Output: $\boldsymbol{X}^* \leftarrow \boldsymbol{X}^k$

TABLE I
NECESSARY PARAMETERS OF TESTED VIDEO

Video Title	Films	News	Star wars	Toy Story	Ball
BL max rate(kbps)	652	739	490	1084	971
EL max rate(kbps)	2000	2000	2000	2000	2000
BL average PSNR	36.7	37.0	37.5	36.6	37.0
Video arrival ratio	0.367	0.221	0.165	0.134	0.113
Mean of β	3.73	3.66	3.39	3.83	3.65

The $\boldsymbol{X}$-update is a proximal operator evaluation, which can be solved by any standard method, such as Newtons method. We use the proximity operator of $g(\boldsymbol{Z})$ to update $\boldsymbol{Z}$.

IV. SIMULATION

In this section we will present our simulation results and point out the key insights observed. Three schemes called QABA [9], UBA and the proposed HQOBA are compared in this simulation. UBA is a streaming scheme that allocate equal bandwidth to each channel. It is assumed that FB divides BL stream and EL stream into $2^5 - 1$ equal-sized segments. Thus they are both broadcasted on five channels. Note that enough bandwidth would be allocated to BL streams to ensure the complete transmission of BL. The full encoding rate of EL is 2000 kbps. The access probability of all videos subjects to a Zipf distribution [17] with the skew parameter θ is equal to 0.27. Then we can define the access probability of video i as $\alpha_i = \frac{1}{i^{1-\theta}} / \sum_{j=1}^{V} \frac{1}{j^{1-\theta}}$. Each stream is accessed by a separate set of five client types. In this simulation, we assume that the available bandwidth changes via CBR (constant-bit-rate) mode. It can be easily extended to the VBR (variable-bit-rate) mode. The other necessary parameters are listed in table I.

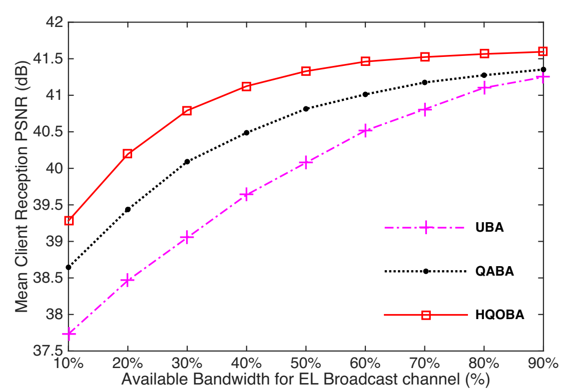
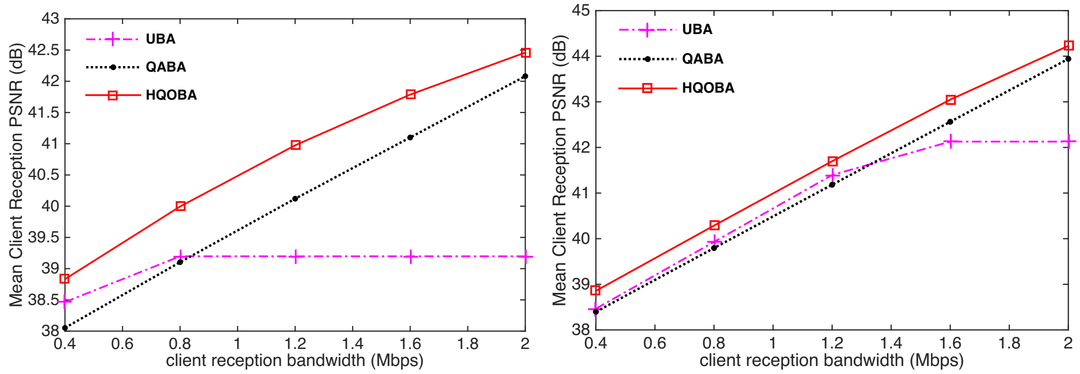


Fig. 4. Mean client reception PSNR obtained by all schemes when the available bandwidth for EL broadcast channels changes

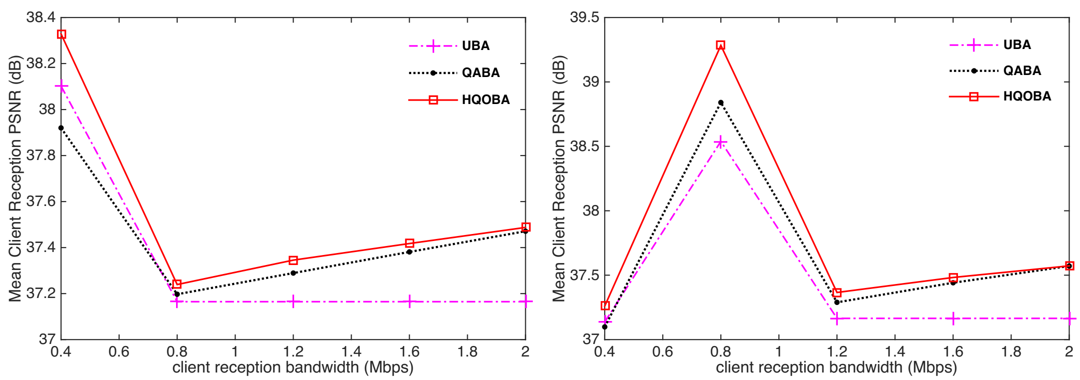
We employ a commonly used metric, namely PSNR (Peak Signal to Noise Ratio) [18], to measure the quality of the video segment client received. Fig. 4 explores the mean client reception PSNR obtained by all schemes for different EL available bandwidth with client reception bandwidth set $[0.4, 0.8, 1.2, 1.6, 2.0]$ $Mbps$. We denote client proportion set as $\mu = [\mu_1, \cdots, \mu_M]$, where μ_m represents the proportion of the type m clients. Given $\mu = [0.2, 0.2, 0.2, 0.2, 0.2]$, we



(a) available bandwidth for EL broadcast channels: 40%

(b) available bandwidth for EL broadcast channels: 70%

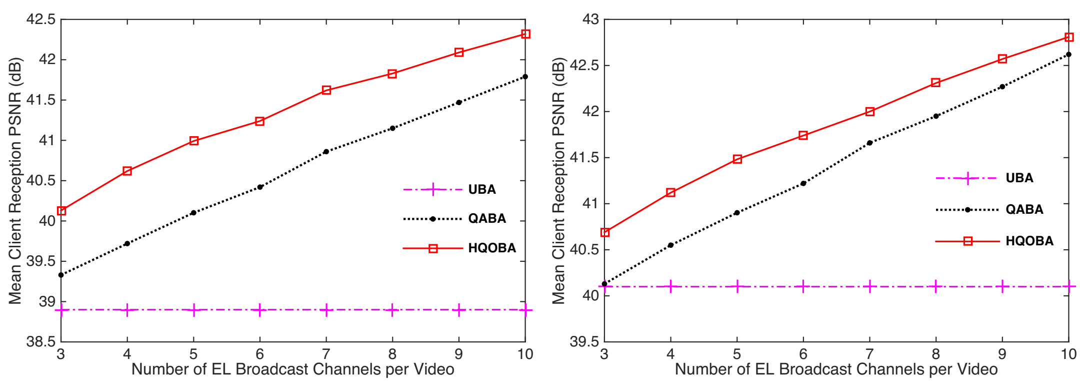
Fig. 5. Mean client reception PSNR comparison under different available bandwidth for EL broadcast channels.



(a) client proportion set: [0.6, 0.1, 0.1, 0.1, 0.1]

(b) client proportion set: [0.1, 0.6, 0.1, 0.1, 0.1]

Fig. 6. Mean client reception PSNR comparison under different client proportion set μ.



(a) available bandwidth for EL broadcast channels: 30%

(b) available bandwidth for EL broadcast channels: 70%

Fig. 7. Mean client reception PSNR obtained by all schemes for different number of broadcast channels

observe that HQOBA yields the largest mean client reception PSNR with different available bandwidth. Since HQOBA takes the heterogeneity of client reception bandwidth into consideration, it does not waste server bandwidth when the client reception bandwidth are low. An interesting perspective is the three curves get closer with the increase of the available bandwidth for EL broadcast channels. This is because the overall improvement of available bandwidth enables all channels transmit at a larger bit-rate.

Specifically, with various client reception bandwidth, we compare the three schemes under different available bandwidth conditions in Fig. 5. The results show that the mean user reception PSNR curve of HQOBA lies entirely above QABA and UBA for each type client in different available bandwidth

conditions. Note that the increase of available bandwidth narrows the gap of the three curves, which also satisfies the results in Fig. 4. An interesting perspective is the trend of PSNR curve of UBA. It achieves PSNR peak at $0.8Mbps$ and $1.6Mbps$ when the available bandwidth for EL broadcast channels are 40% and 70%, respectively. Moreover, it performs better than QABA before the peak PSNR. We can deduce that almost all the client reception bandwidth are lower than the bandwidth allocated in each channel with 40% EL available bandwidth occasionally.

Fig. 6 shows the mean PSNR of each client type under different client proportion sets. The set of client reception bandwidth and available server bandwidth remain the same. In this figure, two client proportion set μ are assigned, namely: a) the client type with bandwidth $0.4Mbps$ is assigned proportion of 0.6 and remaining types are assigned with uniform proportion of 0.1. b) just like (a) but stronger client type with bandwidth $0.8Mbps$ is assigned proportion of 0.6. The results in Fig. 6 (a) indicate that weak clients with reception bandwidth $0.4Mbps$ get preferential treatment over high bandwidth clients by all schemes. The reason of this property is that low bandwidth clients dominate the objective function of these schemes. Among these schemes, HQOBA provides the highest PSNR for dominating clients. Fig. 6 (b) shows that HQOBA still has an advantage when dominating clients reception bandwidth increase to 0.8 $Mbps$. If we assign larger bandwidth client type more weight, the gap between HQOBA and QABA will narrow, which is consistent with the observed result of Fig. 4.

Fig. 7 shows mean client reception PSNR obtained by all schemes when different number of broadcast channels of EL and BL stream is applied. In this simulation, the client reception bandwidth set is $[0.4, 0.8, 1.2, 1.6, 2.0]Mbps$ and client proportion set $\mu = [0.2, 0.2, 0.2, 0.2, 0.2]$. In this figure, two available bandwidth for EL broadcast channels are assigned, namely a) 30%. b) 70%. The results show that HQOBA yields the best performance. It can be observed from Fig. 7 (a) that the visual quality of HQOBA and QABA improves with the increasing of EL broadcast channel number, while UBA keeps the same value all the time. The reason is that HQOBA and QABA takes broadcasting frequency into account and the gap between broadcasting frequencies becomes large with the increase of the number of broadcast channels. It can be observed from Fig. 7 (b) that when available bandwidth for EL broadcast channels increases, the difference between mean client reception PSNR of all schemes decreases, which also demonstrates the rationality of Fig. 4.

V. Conclusion

In this paper, we propose an HQOBA (Heterogeneous Quality-Oriented Bandwidth Allocation) method to address the challenge of VoD streaming in wireless network. On the one hand, we extend the scalability of VoD service with the employment of FGS video coding and Periodic Broadcasting to handle the problem of scarce server bandwidth and the time-varying wireless channel condition in wireless network. On the other hand, in order to tackle the heterogeneous client bandwidth problem, the practical minimization scheme is proposed to optimize reception visual quality for clients equipped with different reception bandwidth capacities. Furthermore, a bandwidth allocation algorithm is developed to maximize client perceptual visual quality for the proposed transmit scheme. The simulation results demonstrate the advantages of HQOBA over QABA and UBA under different server bandwidth conditions, various sets of client reception bandwidth and diverse proportion of each client type.

References

[1] C. V. N. I. Cisco, "Global mobile data traffic forecast update, 2013–2018," *white paper*, 2014.

[2] J. Choi, A. Reaz, and B. Mukherjee, "A survey of user behavior in vod service and bandwidth-saving multicast streaming schemes," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 1, pp. 156–169, 2012.

[3] F. Thouin and M. Coates, "Video-on-demand networks: design approaches and future challenges," *IEEE network*, vol. 21, no. 2, pp. 42–48, 2007.

[4] A. Hu, "Video-on-demand broadcasting protocols: A comprehensive study," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1. IEEE, 2001, pp. 508–517.

[5] K. A. Hua, Y. Cai, and S. Sheu, "Leverage client bandwidth to improve service latency in a periodic broadcast environment," *Journal of Applied Systems Studies*, vol. 2, no. 3, pp. 848–856, 2001.

[6] A. Natarajan, Y. Cai, and J. Wong, "An enhanced client-centric approach for efficient video broadcast," *Multimedia Tools and Applications*, vol. 43, no. 2, pp. 179–193, 2009.

[7] C.-T. Lin and J.-W. Ding, "Car: A low latency video-on-demand broadcasting scheme for heterogeneous receivers," *IEEE transactions on broadcasting*, vol. 52, no. 3, pp. 336–349, 2006.

[8] X. Wang, Z. Zhong *et al.*, "Dere: A buffer saving and controllable video-on-demand broadcasting scheme for heterogeneous receivers," 2015.

[9] J.-W. Ding, D.-J. Deng, T.-Y. Wu, and H.-H. Chen, "Quality-aware bandwidth allocation for scalable on-demand streaming in wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 3, pp. 366–376, 2010.

[10] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h. 264/avc standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1103–1120, 2007.

[11] P. Gill, L. Shi, A. Mahanti, Z. Li, and D. L. Eager, "Scalable on-demand media streaming for heterogeneous clients," *ACM TOMM*, vol. 5, no. 1, p. 8, 2008.

[12] L.-S. Juhn and L.-M. Tseng, "Fast data broadcasting and receiving scheme for popular video service," *Broadcasting, IEEE Transactions on*, vol. 44, no. 1, pp. 100–105, 1998.

[13] D. Marpe, T. Wiegand, and S. Gordon, "H. 264/mpeg4-avc fidelity range extensions: Tools, profiles, performance, and application areas," in *IEEE International Conference on Image Processing 2005*, vol. 1. IEEE, 2005, pp. I–593.

[14] P. Seeling, P. De Cuetos, and M. Reisslein, "Fine granularity scalable video: implications for streaming and a trace-based evaluation methodology," *IEEE Communications Magazine*, vol. 43, no. 4, pp. 138–142, 2005.

[15] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[16] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming," 2008.

[17] L. A. Adamic and B. A. Huberman, "Zipfs law and the internet," *Glottometrics*, vol. 3, no. 1, pp. 143–150, 2002.

[18] S. Winkler, "Video quality measurement standards-current status and trends," *Proceedings of ICICS 2009*, 2009.